Postal Regulatory Commission
Submitted 7/10/2015 8:00:00 AM
Filing ID: 92750
Accepted 7/10/2015

**BEFORE THE**
**POSTAL REGULATORY COMMISSION**
**WASHINGTON, D.C. 20268-0001**

| | | |
|---|---|---|
| PERIODIC REPORTING<br>(PROPOSAL THIRTEEN) | )<br>) | Docket No. RM2015-7 |

# DECLARATION OF CHRISTIAN T. LUNDBLAD ON BEHALF OF AMAZON FULFILLMENT SERVICES, INC.

1. My name is Christian T. Lundblad. I submit this Statement on behalf of Amazon Fulfillment Services, Inc.

## I. AUTOBIOGRAPHICAL SKETCH

2. I am the Edward M. O'Herron Distinguished Scholar and Professor of Finance at the University of North Carolina's Kenan-Flagler Business School. I have been on the faculty of the school since 2006. I was Assistant Professor of Finance at Indiana University from 2001-2006. During 2000-2001, I served as a financial economist at the Federal Reserve Board in Washington, D.C. I am an Associate Editor for the Journal of Finance, the leading academic finance journal.

3. I have a Ph.D. in financial economics and an M.A. in economics from Duke University, and a B.A. in economics and English literature (with highest honors) from Washington University in St. Louis.

4. My research spans asset pricing and international finance, with a specialization in emerging market development, and a heavy reliance on time-series

econometrics. My research has been published in leading academic journals such as the Journal of Finance, the Review of Financial Studies, and the Journal of Financial Economics, and has been cited in general press by outlets such as The Economist and Reuters. I previously appeared before the Postal Regulatory Commission in Docket No. R2013-11, in which I submitted a written statement for a coalition of mailers on the effect of the 2007-2009 recession on postal finances.[1]

5. Of particular relevance to this proceeding, my published research,[2] while in a different context, uses many of the same econometric methods (including panel and cross-sectional regressions). While addressing financial market questions, several of these research papers explore a comparable relationship between price changes, as the dependent variable, and quantity or volume, as the explanatory variable. My curriculum vitae is attached to this Statement.

## II. SUMMARY

6. Amazon has asked me to review and comment on the June 8, 2015 Supplemental Report of Dr. Kevin Neels on behalf of United Parcel Service ("UPS") and related portions of the UPS comments to which the statement is attached.[3] As detailed below, I believe that both of the models proposed by Dr. Neels are too deeply flawed for

[1] Docket No. R2013-11, Statement of Christian T. Lundblad, November 26, 2013.

[2] See, among others, Does Financial Liberalization Spur Growth (2005) *Journal of Financial Economics* for a careful use of panel regressions and Regulatory Pressure and Fire Sales in the Corporate Bond Market (2011) *Journal of Financial Economics* for quantity-price regressions.

[3] Docket No. RM2015-7, United Parcel Service Comments Attaching Supplemental Report Related to Proposal Thirteen (filed June 8, 2015) ("UPS Comments").

the Commission to accept. By contrast, the criticisms of the Postal Service's model offered by Dr. Neels and UPS lack real-world significance.

7. In Section III below, I discuss the purpose of the competing econometric models. In Section IV, I describe the two models advocated by Dr. Neels in his June 8 Supplemental Report and identify the most important flaws in those models:

- The main flaw in Dr. Neels' preferred model, the "National Form 3999 Model," is its reliance on imputed values for deviation parcel, in-receptacle parcel, and collection mail volumes for use in his regressions because no actual data are available for these variables for the vast majority of ZIP Codes in his data set. Dr. Neels Supplemental Report at 27. His imputation approach creates substantial multicollinearity, meaning that the explanatory variables in this model are highly correlated. This multicollinearity prevents the model from accurately estimating the independent effect of individual variables (e.g., parcel volumes) on city carrier street costs. Thus, the model cannot accurately estimate the marginal cost of each type of mail. In addition, this model is not robust to changes in the specification of the imputation regressions; that is, changes in specification result in substantial changes in results. This further reduces the confidence that the Commission can have in this model.

- The primary flaw in Dr. Neels' fallback model, "Modified Proposal 13,"[4] is its reliance on Form 3999 parcel volume data[5] that all parties (including UPS) and all of their experts (including Dr. Neels) agree are of very low quality. Although poor data quality sometimes can simply cause the effect of an explanatory variable in a regression to be understated, this is unlikely to be the case here. Additionally, the Form 3999 parcel variables that Dr. Neels inserts into the model are highly correlated with other explanatory variables. This multicollinearity makes it difficult to disentangle the independent effect, if any, of parcel volumes on regular delivery costs.

8. In Section V, I explain why the criticisms of the Postal Service's proposed model offered by Dr. Neels and UPS are detached from real-world economic practice. Econometricians always prefer to have access to a perfect data set—e.g., of perfect data quality, collected with the exact definitions needed for modeling purposes—collected for other (e.g., operational) purposes and thus available to the analyst at no additional cost, but this perfection is almost always unattainable.

9. The Postal Service, faced with practical realities, such as the flaws in the Form 3999 parcel volume data and the presence of substantial multicollinearity between relevant volume variables, made reasonable choices – e.g., using operational data where

[4] "Proposal 13" refers to Docket No. RM2015-7, *Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Change in Analytical Principles (Proposal Thirteen)*, initiated on December 11, 2014. The "Modified Proposal 13" was initially presented by Dr. Neels on pages 9-10 of his March 18, 2015 report.

[5] We refer to the parcels measure in the Form 3999 data (which was originally collected in the DOIS dataset) as the Form 3999 parcel volumes, consistent with the terminology used by other parties. See for example, Dr. Neels' Supplemental Report at 28.

appropriate, but filling in data gaps with field studies and separately analyzing regular and parcel delivery time. Filling data gaps with reasonably-designed field studies is far preferable to using data that are clearly flawed.

10. UPS repeatedly criticizes the Postal Service's model for supposedly failing to "let the data speak" on whether parcel volumes affect regular delivery costs. UPS Comments at 2. I agree that where the data can speak accurately, the analyst should listen, but that is not always possible. UPS argues that, through the Modified Proposal 13 Model, the data told Dr. Neels that there is a significant relationship between parcel volumes and regular delivery time. UPS Comments at 19. I disagree. The data flaws and multicollinearity in Dr. Neels' Modified Proposal 13 Model leave it with nothing meaningful to say about the relationship between parcel volumes and regular delivery time.

## III. THE PURPOSE OF THE ECONOMETRIC MODELS IN THIS PROCEEDING.

11. The purpose of the econometric models presented in this proceeding is to estimate the variability of city carrier street costs with respect to mail volume by type (e.g., delivery point sequenced letters, cased mail, sequenced mail, in-receptacle parcels, deviation parcels). This modeling is an intermediate step in the method for attributing city carrier street costs to individual products.

12. Given this purpose, an acceptable model must be able to identify the independent effect of each of these types of mail on city carrier street costs. A model that can explain variations in city carrier street costs, but cannot identify the individual effect of each variable, is insufficient.

13. For this reason, multicollinearity – a high degree of correlation between two or more explanatory variables – can limit the ability to estimate the effect of individual explanatory variables on the dependent variable. This is a critical concern, and much, but not all, of my review of the UPS-proposed models focuses on this issue.

## IV. NEITHER OF UPS'S PROPOSED COSTING METHODS ACCURATELY ESTIMATES THE RELATIONSHIP BETWEEN VOLUME AND CITY CARRIER STREET TIME.

### A. Description Of The Two Models Sponsored By Dr. Neels In His June 8 Supplemental Report

14. In his June 8 statement, Dr. Neels advances two econometric models for estimating this relationship.

15. His preferred model – the **National Form 3999 Model** – is a two-stage model. In the first stage, Dr. Neels performs three separate regressions to impute in-receptacle parcel, deviation parcel, and collection volumes for approximately 11,000 ZIP Codes served by city carriers based upon actual volume data collected by USPS in its field studies of 300 ZIP Codes. Neels Supplemental Report at 27.

16. The second stage of this model is a cross-sectional regression, run across ZIP Codes, using a quadratic functional form similar to that used in the Postal Service's own model. Dr. Neels Supplemental Report at 27. The dependent variable in the second-stage regression is total street time. The explanatory variables are derived from (1) the imputed volumes from the first stage, (2) Form 3999 volumes for other types of mail, and (3) other ZIP Code characteristics. Neels Supplemental Report at 45, Table 16.

17. Dr. Neels' fallback proposal – the **Modified Proposal 13 Model** – relies on a variant of the Postal Service's regular delivery model, to which he adds Form 3999 parcel volumes as an explanatory variable. Based on the results of this model, Dr. Neels recommends distributing 2.9 percent of regular delivery costs to individual products using a parcel distribution key. UPS Comments at 8, 18; Neels Supplemental Report at 43.

### B. Dr. Neels' reliance on imputed parcel and collection mail volumes renders the National Form 3999 Model useless for costing.

18. The primary flaw in the National Form 3999 Model is its use of imputed parcel and collection mail volumes. Substantial multicollinearity results, rendering an assessment of marginal costs impossible.

19. First, on the imputation of data, the Form 3999 data set lacks some of the volumes necessary to directly measure the relationship between city carrier street costs and mail volume (by type). Specifically, the Form 3999 data set lacks deviation parcel, in-receptacle parcel, and collection volumes. The Postal Service solves this problem by performing field studies of 300 ZIP Codes to collect this information.

20. Dr. Neels makes a different choice. Using volume data for the 300 ZIP Codes for which USPS collected parcel and collection mail volume, he attempts to impute values for these three volume variables for the remaining ZIP Codes. Specifically, he fits three binomial regressions (one for each of these three types of mail) onto several candidate explanatory variables that are available for the larger collection of ZIP Codes in the Form 3999 data set.

21. When imputing missing data, the analyst hopes that the fits of the regressions will be very tight, in which case one may be able to use the imputed data to extend reliably the second-stage model to the full cross-section of ZIP Codes. This is not so here. The fits of his imputation regressions are not tight, and his choices of explanatory variables in the imputation regressions create substantial multicollinearity in his second-stage model.

22. The three first stage regressions are measured with relatively low $R^2$s. Simply put, the data we possess in the Form 3999 data set for the larger collection of ZIP Codes do not explain well the variability in the three missing volumes of interest. For example, the $R^2$ of the special field study deviation parcel imputation is only 54 percent, meaning approximately half of the variability of the field study deviation parcel volume is left unexplained. Similarly, the $R^2$'s for the separate regressions of in-receptacle parcel and collection volume field study data are only 63 percent and 31 percent, respectively. Neels Supplemental Report at 32.

23. An important side effect of the relatively low $R^2$'s in the first stage is that the use of the imputed (with error) variables in the second stage regression increases (and likely increases significantly) the standard errors on the relevant coefficients of interest. That is, we are essentially left with what is commonly called an "errors-in-variables" problem. The deviation parcel, in-receptacle parcel, and collection mail volumes that are used in the second-stage regression are not directly observed, but imputed with considerable error from the first stage exercise. Hence, the standard errors on the second stage estimates are significantly underreported.

24. Dr. Neels acknowledges this point in passing (Neels Supplemental Report at 38) but fails to acknowledge its true significance. The effect on the relevant standard errors associated with the coefficients – serving as critical ingredients in the costing analysis – is likely quite sizeable. This leaves our inferences about any marginal cost figure considerably less precise than suggested by Dr. Neels' reported figures (which include no allowance for this first-stage estimation error). This alone should give the Commission considerable reason for pause. However, the method suffers from an additional, even larger concern that renders the exercise entirely without merit.

25. The second (far larger) concern revolves around the specific variables that Dr. Neels employs to conduct the imputation. Specifically, his use of several explanatory variables from the second-stage regression – including the number of delivery points and mail volumes of other shapes (e.g., delivery point sequenced mail) (Neels Supplemental Report at 27) – as explanatory variables in the first stage regression is highly problematic. This sharing of explanatory variables between the first and second stage regressions results mechanically in a sizeable degree of multicollinearity between volume variables in the second stage costing regression, rendering it quite difficult to disentangle the effects of individual explanatory variables on city carrier street time.

26. As mentioned above, multicollinearity is a situation in which there is a very high degree of correlation between the independent explanatory variables in a multiple regression. When multicollinearity is present, even modest changes in the data can produce wide swings in the resulting coefficient estimates, wreaking havoc on the analysis and limiting the conclusions one can draw from it.

27. To provide a simple example, suppose that one tried to impute the volumes of deviation parcel, in-receptacle parcel, and collection mail by regressing the three special field study volume series on the number of delivery points alone (again, three separate regressions). The regression coefficient would be statistically significant in each case, and one might conclude that the number of delivery points is a good predictor of these volume series more broadly (in instances when they are not measured).

28. The imputed data for each of the separate missing volumes would now look like ($a_i + b_i * DP$), where *i* is for one of the three imputed series. One would then conduct the second-stage regression on the measured volume series, these three imputed series, *DP* directly, and a collection of other variables. This regression would not be estimable because four of the explanatory variables would be perfectly correlated (one is *DP* directly and the other four represent a linear combination of *DP*). In such a case, the multicollinearity is so extreme as to preclude even estimating the model. The regression would teach us nothing about the true economic effects of the three volume series, *DP* itself, and any other variable. Unfortunately, this simple example is close to the reality of Dr. Neels' National Form 3999 Model.

29. The correlations between the three imputed volumes from Dr. Neels' preferred imputation method and DP are 93 percent for deviation parcels, 90 percent for in-receptacle parcels, and 66 percent for collection mail, respectively. This substantial correlation is illustrated in Figures 1 through 3 below.

**Figure 1**

**Scatter Plot of Imputed Deviation Parcel Volume vs. Delivery Points**

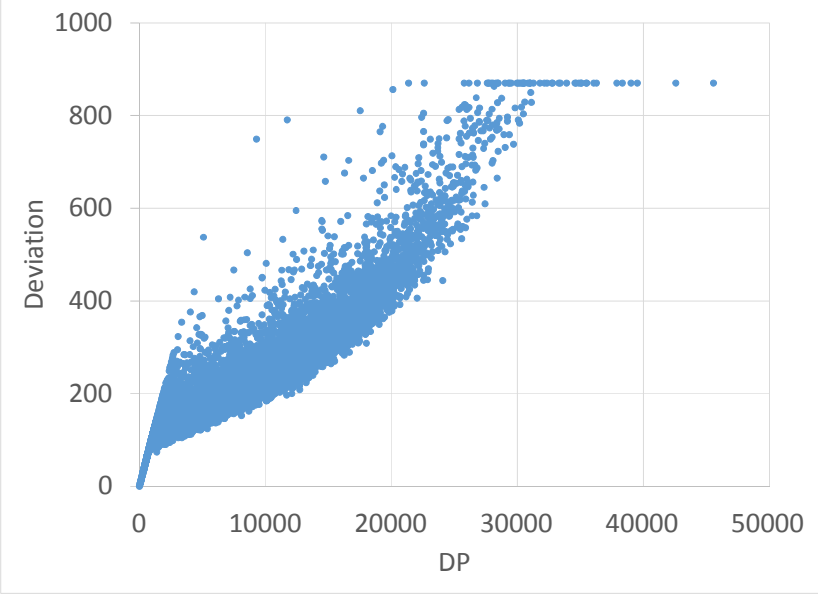



**Figure 2**

**Scatter Plot of In-Receptacle Parcel Volume vs. Delivery Points**

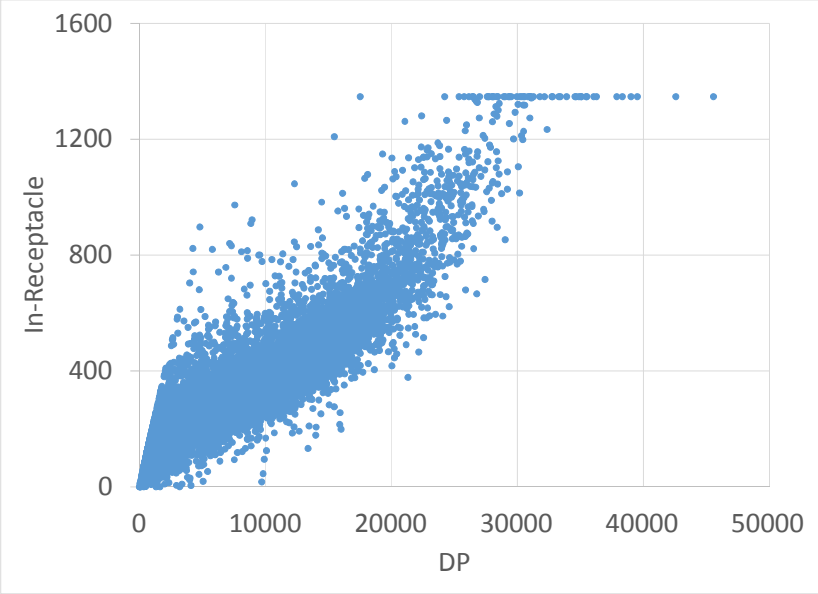
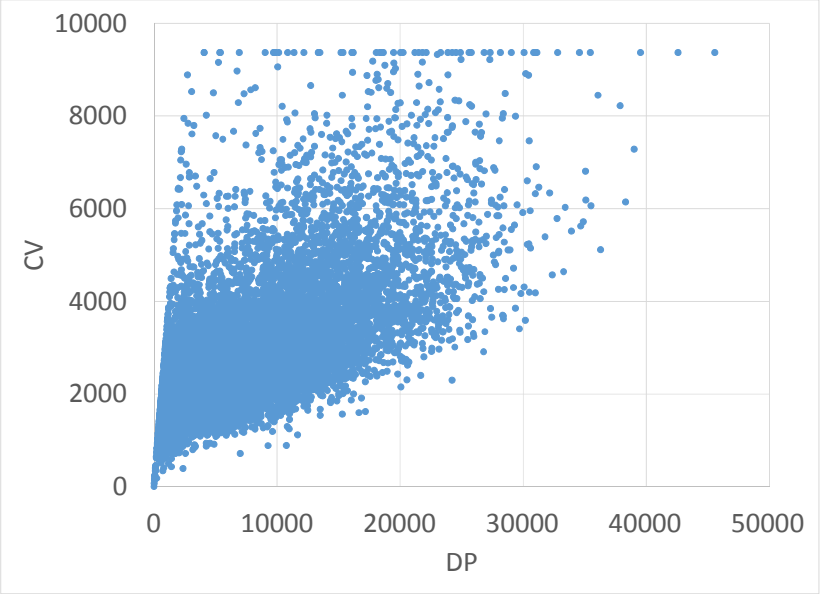


**Figure 3**

**Scatter Plot of Collection Volume vs. Delivery Points**




30. That is, the three imputed volumes (particularly the parcel volumes) that Dr. Neels employs in the second stage costing regression are largely determined by the same explanatory variable (and one that is already an explanatory variable in its own right in the same regression).[6]

31. I now turn to the degree to which the second stage costing regression is plagued by multicollinearity. The estimates and standard errors provided in Dr. Neels' report (Neels Supplemental Report at 13, Table 1) are skewed by their failure to include an adjustment for the first-stage imputation error. Here, for each coefficient, I provided an additional statistic, the variance inflation factor ("VIF"), a term that is designed to measure the extent to which the regression is plagued by multicollinearity.

[6] Interestingly, and in sharp contrast, the parcel volumes actually recorded in Form 3999 play a relatively small role in imputing deviation parcel volume. I discuss the implications of this in ¶ 48, below.

32. As noted by the Postal Service on page 70 of its December 11, 2014 initial report on the Postal Service's City Carrier Street Time Study, this statistic is calculated as follows: The researcher runs *N* different auxiliary regressions, one for each explanatory variable in the main model of interest. The auxiliary regressions are estimated, variable-by-variable, by regressing the particular explanatory variable *i* onto the other explanatory variables in the main model. Then, for each explanatory variable *i*, the $VIF_i$ is equal to $(1 / (1- R_i^2))$ where $R_i^2$ is from the regression of explanatory variable *i* on all the other explanatory variables. This is, in essence, telling you how closely related the explanatory variables are to one another. If variable *i* is completely unrelated to the other explanatory variables (i.e., no multicollinearity exists), then the $R_i^2$ will be zero and $VIF_i$ will be 1. If multicollinearity associated with this variable is problematic, the $R_i^2$ will be quite high (perhaps approaching 1), and the VIF can be very large. A rule of thumb is that if the VIF for a coefficient exceeds 10, then multicollinearity associated with that variable is high. Such a threshold is only exceeded when the $R_i^2$ (from the regression of explanatory variable *i* on the other explanatory variables) is larger than 90 percent.

33. Table 1 below reports the VIFs for each coefficient in the second stage of Dr. Neels' National Form 3999 Model. As you can see, only 6 of the more than 60 explanatory variables used in this regression are not associated with significant multicollinearity problems. Further, nearly every variable from the set of imputed volumes is associated with significant multicollinearity concerns with VIFs an order of magnitude or two above the rule of thumb of ten.

**Table 1**

**VIFs of Dr. Neels' 2nd Stage Model**

| Coefficent | VIF | Coefficent | VIF |
|---|---|---|---|
| DPS | 227.37 | CM*CV | 96.16 |
| DPS2 | 806.14 | CM*IRP | 413.09 |
| DPS*FSS | 160.09 | CM*DP | 372.62 |
| DPS*SEQ | 171.7 | CM*DM | 35.84 |
| DPS*CM | 500.44 | CM*SMDP | 781.51 |
| DPS*Dev | 2696.83 | Dev | 238.51 |
| DPS*CV | 367.51 | Dev2 | 963.17 |
| DPS*IRP | 1795.34 | Dev*CV | 452.99 |
| DPS*DP | 1503.93 | Dev*IRP | 2622.66 |
| DPS*DM | 94.62 | Dev*DP | 1940.85 |
| DPS*SMDP | 1310.33 | Dev*DM | 119.56 |
| FSS | 40.91 | Dev*SMDP | 29.48 |
| FSS2 | 26.4 | CV | 31.74 |
| FSS*SEQ | 12.81 | CV2 | 34.68 |
| FSS*CM | 24.81 | CV*SMDP | 10.37 |
| FSS*Dev | 195.06 | CV*IRP | 326.92 |
| FSS*CV | 31.37 | CV*DP | 216.1 |
| FSS*IRP | 142.58 | CV*DM | 13.88 |
| FSS*DP | 98.9 | IRP | 124.75 |
| FSS*DM | 8.11 | IRP2 | 764.31 |
| FSS*SMDP | 3.96 | IRP*DP | 1159.49 |
| SEQ | 28.59 | IRP*DM | 74.84 |
| SEQ2 | 20.23 | IRP*SMDP | 7.68 |
| SEQ*CM | 37.28 | DP | 129.47 |
| SEQ*Dev | 334.31 | DP2 | 652.73 |
| SEQ*CV | 37.64 | DP*DM | 54.51 |
| SEQ*IRP | 201.83 | DP*SMDP | 21.92 |
| SEQ*DP | 153.06 | DM | 30.59 |
| SEQ*DM | 7.13 | DM2 | 20.77 |
| SEQ*SMDP | 2.13 | DM*SMDP | 4.02 |
| CM | 78.52 | SMDP | 8966.08 |
| CM2 | 83.47 | SMDP2 | 11677.75 |
| CM*Dev | 652.82 | | |

34. This muticollinearity stems from two sources. One is the induced correlation that arises from the fact that the imputed volume measures share many of the same explanatory variables with each other as well as with variables that also directly appear

in the second-stage costing regression. Second, unlike the model presented by USPS, in which multicollinearity is carefully addressed,[7] Dr. Neels allows all of the squared and cross-product terms to appear in the regression, with no regard for the high degree of correlation many of those explanatory variables naturally have among themselves. The presence of this substantial multicollinearity precludes any meaningful inference about relative coefficient magnitudes (let alone any reliable estimates of marginal cost, the ultimate goal of the exercise).

35. To illustrate the effect of Dr. Neels' imputation approach on the results of the National Form 3999 Model, I demonstrate the high degree of sensitivity that the multicollinearity engenders by considering a few other reasonable alternatives for the set of explanatory variables that feature in his first-stage imputation. To proceed, I consider four sets of imputation methods (in addition to again reporting Dr. Neels' imputation for comparison purposes):

1. Full Imputation Model: All of Dr. Neels' candidate explanatory variables are used to impute the three missing volume series;

2. Exclude Mail Volume Variables because they directly feature in the second stage model directly;

3. Exclude Mail Volume and Delivery Point Variables because they directly feature in the second stage model; and

[7] USPS Report on the City Carrier Street Time Study (December 11, 2014) at 68-75.

4. Exclude Mail Volume, Delivery Point, and other explanatory variables that are not statistically significant.

36. As you can see in Table 2, the imputation $R^2$'s are highly sensitive to the choice of explanatory variables used in the imputation model. Clearly, the most important explanatory variable seems to be *DP*; the $R^2$'s decline significantly when *DP* is not used.

**Table 2**

**$R^2$'s of 1st Stage Regressions**

| | Deviation | In-Receptacle | Collection Volume |
|---|---|---|---|
| Neels' Model | 54% | 63% | 32% |
| Full Model | 70% | 67% | 47% |
| Excluding Mail | 67% | 66% | 43% |
| Excluding Mail, DP | 41% | 32% | 25% |
| Excluding Mail, DP, Non-Significant | 35% | 25% | 17% |

37. I view all of these choices as equally plausible because imputation is necessarily a relatively unguided statistical exercise. That said, the versions that do not include variables that also feature directly as explanatory variables in the second stage are more natural because those versions do not mechanically induce multicollinearity.

38. The important lesson from this, however, is the extent to which the choice of imputation model affects the results of the second-stage costing regression. Interestingly, the second stage $R^2$ is largely unaffected by this choice (see Table 3 below). This is unsurprising when multicollinearity is present. The variability of the dependent variable (city carrier street time) that is explained is largely unchanged because the

symptom of multicollinearity is not a low $R^2$, but rather substantial sensitivity with respect to the attribution of where the explained variability is coming from.

**Table 3**

**$R^2$'s and Marginal Costs of 2nd Stage Regressions**

| | Neels' Model | Full Model | Excluding Mail | Excluding Mail, DP | Excluding Mail, DP, Non-Significant |
|---|---|---|---|---|---|
| $R^2$ | 95% | 95% | 95% | 95% | 95% |
| Dev Varability | 12% | -6% | -9% | -11% | -8% |
| CV Varability | 5% | 4% | 4% | 5% | 4% |
| IRP Varability | 4% | 6% | 11% | 9% | 3% |
| Total Varability | 56% | 50% | 46% | 45% | 42% |
| Dev MC | 162.1 | -81.25 | -124.9 | -135.1 | -91.89 |
| CV MC | 5.745 | 9.644 | 3.938 | 4.514 | 4.025 |
| IRP MC | 38.42 | 47.08 | 96.47 | 69.48 | 25.12 |

39. Here, one can see five plausible imputation exercises generating five very different impressions of the marginal cost of delivering deviation parcels, in particular. Indeed, the cost results are quite sensitive to the specification of the first-stage imputation; we actually uncover an odd negative marginal cost for deviation parcel volume in some cases. Taken together, all one can reliably conclude is that the Dr. Neels' National Form 3999 Model suffers from a problem that renders economic inference about relevant costs impossible. The model should be rejected.

### C. Multicollinearity and poor data quality prevent Dr. Neels' Modified Proposal 13 Model from establishing any reliable relationship between parcel volume and regular delivery time.

40. As a fallback in case the Commission rejects Dr. Neels' "National Form 3999 Model," he proposes in the alternative that the Commission modify Proposal Thirteen to include parcels in the equation for regulator delivery time. The Commission should decline to adopt this model as well.

41. Dr. Neels' fallback model also suffers from a severe multicollinearity problem. Because of this, I would be very uncomfortable concluding from this model that there is a relationship between parcel volumes and regular delivery time. Table 4 below provides the VIFs for each variable in this model involving the Form 3999 parcel variable:

**Table 4**

**VIFs for Each Variable**

| Variable | VIF |
|---|---|
| par | 26.0 |
| par2 | 11.4 |
| pardps | 33.2 |
| parcm | 14.1 |
| parseq | 5.8 |
| parfss | 7.3 |
| parcv | 11.1 |
| parpd | 34.4 |

42. As Table 4 shows, 6 of the 8 explanatory variables for the Form 3999 parcel volume (its linear, square, and cross-product terms) suffer from multicollinearity. The size of the VIFs for the parcel variables indicates that the coefficients on the parcel volume variables are picking up correlations to other variables, not a real relationship between parcel volume and regular delivery time. For example, the parcels variable ("par") has a correlation greater than 30 percent with 19 of the 35 variables, As with the National Form 3999 Model, one simply cannot reliably infer relevant marginal costs (as derived from regression coefficients) from a regression model with such severe multicollinearity.

43. Apart from multicollinearity, the Modified Proposal 13 Model suffers from data problems of the first order. Indeed, accurate Form 3999 parcel volume data simply are not available, making Dr. Neels' model nothing more than illustrative of what could be done at some future point in time if higher-quality data were to become available. Even if the Commission were inclined to accept this model when data quality improves, current data litigations argue against doing so now.

44. The Form 3999 parcel volume data indisputably have limitations; the USPS was the first to acknowledge this:

> Although the Postal Service explained to UPS, as detailed in the Postal Service's Notice of the filing of this material, that it does not view the requested data as comparably suited for analysis as the other DOIS data used in the Postal Service analysis, the Postal Service nonetheless in USPS-RM2015-7/2 provides the requested data….The PARCELS variable contains an operational count by route of 'large' parcels for those routes on those days in which information appeared in the DOIS dataset. Unlike other volumes which are done by machine counts or linear measurements, accurate parcel counts are cumbersome to complete and may not be done, which accounts for the large amount of zero values in the data.

USPS Library Reference USPS-RM2015-7/2, Preface (filed March 2, 2015).

45. UPS has agreed. UPS Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs (March 18, 2015) at 27. So has Dr. Neels:

> Inspection of the DOIS parcel data suggested that the Postal Service's concern about data quality were not groundless. The data set did contain large numbers of zero values, and moreover, there appeared to be a pattern to the distribution of these zero values.

Neels Supplemental Report (March 18, 2015) at 7-8.

46. In subsequent filings, UPS again noted the deficiencies in the Postal Service's data. UPS Comments (June 8, 2015) at 5 & 18. I have reviewed these parcel volume data and agree that they are problematic.

47. In particular, it is important to emphasize that, given the Postal Service's caveats about data quality, one would expect that the data problems are not limited to values of zero. A comparison of deviation parcel volumes from the special field study to Form 3999 parcel volumes confirms this expectation. For the 300 ZIP Codes for which both deviation parcel and Form 3999 parcel volume data are available, we can explore the degree to which these two variables correspond to one another. Indeed, this is a central building block of Dr. Neels' imputation method in the National Form 3999 Model; Form 3999 parcel volumes do feature as an explanatory variable in his imputations.

48. In Dr. Neels' imputation, very little of the variation in deviation parcel volume is explained by his Form 3999 parcel volume per delivery point measure (as previously noted, the imputation is driven largely by the DP (delivery points) variable). The low explanatory power of this volume measure strongly suggests the presence of issues that go beyond missing volumes. Additionally, although running the first-stage binomial regression of deviation parcel volume on Dr. Neels' Form 3999 parcel volume per delivery point measure alone generates a statistically significant coefficient for parcel volume, the $R^2$ of that regression is only 7.55 percent. Running the regression on Form 3999 parcel volumes, rather than Neels' per-delivery point measure, produces a higher, but still low $R^2$ of 30.6 percent.

49. Dr. Neels dismisses this concern by arguing that, when an explanatory variable is contaminated by random noise, the relevant regression coefficient is biased

downward. This "attenuation bias," he suggests, means that any estimation error is likely to have *understated* the correlations shown by this model, and the cost attributions implied by them. Neels Supplemental Report at 43, n. 51. While attenuation bias is a common finding in a linear regression with one explanatory variable,[8] a few features of Dr. Neels' specification complicate this situation.

50. First, the presumption of attenuation bias no longer holds in multivariate linear regression models, and the Modified Proposal 13 model is a multivariate regression. (For a recent review article, see Chen, Hong, and Nekipelov, "Nonlinear Models of Measurement Errors," Journal of Economic Literature 49:4, 901–937 (2011).) Second, when the specification is nonlinear, the bias is also not as clear; the intuition for linear models can potentially lead us astray for nonlinear models. *Id*. Finally, a critical condition for establishing attenuation bias is that the noise characterizing the parcel volume data is indeed entirely random. If the incidence of missing data is correlated with other relevant data employed in the specification, then the estimator is very hard to interpret. *Id*. As Dr. Neels noted in the statement quoted above, there "appeared to be a pattern to the distribution of these zero values." In other words, the noise does not appear to be random.

## V. RESPONSE TO DR. NEELS' CRITICISMS OF USPS COST MODEL

51. Dr. Neels and the UPS also devotes much of their July 8 supplemental comments and Supplemental Report to criticisms of the Postal Service's Proposal 13 Model (the "USPS Model"). In particular, they focus on the Postal Service's: (1) use of a

[8] *See* Ragnar Frisch, Statistical Conference Study. Oslo: University Institute of Economics (1934).

panel data set; (2) assumption that parcel volume is unrelated to regular delivery time; (3) reliance on field studies; and (4) data scrubbing. These criticisms are unfounded, and certainly do not justify adoption of the flawed models of Dr. Neels instead of the USPS Model. I discuss each criticism in turn.

### A. The Postal Service's Use Of A Panel Data Set Was Appropriate.

52. Dr. Neels' criticism that USPS's use of a panel data set produces "short-run" variability estimates is unfounded, and his collapsing of the USPS panel data set (by running cross-sectional regressions on ZIP Code-level averages) destroys relevant information available in the data. Neels Supplemental Report at 10.

53. Generally speaking, panel regression techniques should be employed if we have both time-series (here, within ZIP Code) and cross-sectional (here, across ZIP Code) data from which we can make inferences about the relevant economic relationships (here, the relationship between volume and city carrier street time).

54. The argument against the use of panel data hinges on a statement made by Dr. Neels that the across-ZIP Code data are more informative about "long-run" cost relationships than the within-ZIP Code temporal data. Neels Supplemental Report at 10. First, the notion that cross-sectional variation contains more information about long-run costing relationships is unclear and not well-defined. I infer that the statement is intended to mean that ZIP Code level operations have largely achieved maximal efficiency, whereas temporal variation reflects other concerns. In contrast, throwing away data is ill-advised unless one is quite sure that the data have little to say about the relevant economic relationships one is trying to extract. Before throwing away data, I would at

least evaluate the degree to which the two sources of variation are contributing to the regression output.

55. In evaluating the sources of information that are playing the largest role in identifying the relevant elasticities in the main panel regression in the USPS Report, two main points are important. I ran two versions of the panel regression, examining the separate degrees to which cross-sectional and time-series variation contribute to the full panel regression result. First, the volume measures in the cross section do explain more variation than the time-series dimension. Most of the variables in the cross-section are statistically significant. This would support the idea that cross-ZIP Code variation is central – even in the panel regression proposed by USPS. That fact, however, does not mean the temporal variation is without value. Indeed, two volume measures are significant at the 10 percent level in the time-series dimension, so it is important. To the extent that temporal variation is of any importance, the panel data structure employed within USPS Report is preferred.

### B. Dr. Neels Has Failed To Establish Any Significant Relationship Between Parcel Volume and Regular Delivery Time.

56. Dr. Neels' claim that there is a significant relationship between parcel volume and regular delivery time is also unsupported. To be clear, I do not have the operational expertise to know whether parcel delivery is perfectly separable from letter and flat delivery, and, as an econometrician, I am sympathetic to Dr. Neels' desire to empirically test this relationship. Dr. Neels' model, however, does not prove that there is a relationship. As described earlier, Dr. Neels' fallback Modified Proposal 13 Model

employs low-quality Form 3999 parcel data and suffers from significant multicollinearity problems.

### C. The Postal Service Reasonably Relied On Field Studies.

57. Dr. Neels' criticisms of the Postal Service's reliance on field studies are also unfounded. There is a long history in economics, marketing and related fields of using data collected from field studies (see, for example, the huge and ongoing economics literature that employs the Panel Study of Income Dynamics, a long-running household panel survey). While now quite vast, this study started as a field exercise, collecting information from nearly 5,000 families in 1968. Recent advances in behavioral economics have relied heavily on field studies. See, e.g., Ariely, Bracha, and Meier, "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," *American Economic Review* 99(1): 544-55 (2009). Indeed, a major benefit of field studies is the control that the analyst has over the data collection process, control that analysts generally lack with respect to data collected for other, e.g., operational, purposes.

58. Given the current choice between using data from the Postal Service's field study or using flawed operational data (such as the Form 3999 parcel volume data), the easy choice would be to perform a field study, as the Postal Service did. Even in the future, performing field studies to fill in data gaps may continue to be the best option.

59. It is easy to view, as UPS does, operational data as a panacea on the assumption that operational data will be more comprehensive and less costly for the analyst to obtain. This will be true in some cases, but not others. Specifically, the benefit of using operational data is affected by multiple factors, of which two key ones are whether

(1) operations personnel view the data as being valuable for their purposes; and (2) operations personnel collect the data in a way that is compatible with the analytical framework.

60. If operations personnel do not view the data as valuable for their organization's own purposes, the collection of operational data could become little more than a very high cost field study overseen and implemented by an organization uninterested in the quality of the data. The data collection is likely to be much more costly than a field study – in terms of employee time – due to its comprehensiveness. And if the operations department does not view the data as being valuable for its own purposes, the data are likely to be poor.

61. With respect to the second factor, because operational departments have different uses than economists for data, there is no a priori reason to believe that operational data will be collected in a manner that is compatible with the analytical framework. Thus, the analyst is often left trying to do the equivalent of forcing a square peg into a round hole. Sometimes the peg can be forced into the hole, but this is far from ideal.

62. The current state of Form 3999 parcel volume data provides an informative example of these issues. As discussed elsewhere, the quality of the data is low because "accurate parcel counts are cumbersome to complete"[9] and presumably the value of the data, at least at this point, has not justified the effort to the employees and managers responsible for collecting the data. Additionally, as explained by the Postal Service in

[9] USPS-RM2015-7/2 – Public Materials Filed in Response to Informal Request for Additional DOIS Data, March 2, 2015.

response to an informal UPS question, the Form 3999 definition of a parcel – "1) larger than a shoebox or 2) heavier than two pounds"[10] – is different from those of in-receptacle and deviation parcels used in the city carrier street cost study.

63. Given the issues with the current Form 3999 parcel volume, the Postal Service's decision not to rely on those parcel volume data in its model was entirely understandable. The Commission should carefully scrutinize future updates to these data before approving their use in costing studies.

### D. Dr. Neels' Criticisms Of The Postal Service's Data Scrubbing Are Unfounded.

64. Finally, I disagree with UPS' criticisms of the Postal Service's data scrubbing. UPS Comments (June 8, 2015) at 12. There is also a long history in statistics and econometrics covering the trimming of obvious outliers. See, e.g., S.M.Stigler, "Do Robust Estimators Work with Real Data?", *Annals of Statistics* 5:1055–98 (1977). Data trimming follows the advice of the famous mathematician Legendre, who recommended deleting those observations with errors "too large to be admissible." Adrien-Marie Legendre, "On the Method of Least Squares," in *Nouvelles Méthodes Pour La Détermination des Orbites des Comètes* (1805), translated and reprinted in 2 D.E. Smith, *A Source Book in Mathematics* 576-79 (Dover reprint 1959). That is, if the data are patently wrong or implausible, then dropping them is entirely reasonable. While there are sophisticated econometric techniques available when one is concerned that the relevant data are arbitrarily truncated or censored, that does not seem to be the main issue here.

[10] Notice of the United States Postal Service of Providing Informal Responses to UPS Questions, May 28, 2015, at 2.

Rather, the key issue here is how many data items are you dropping (and why) – and it is useful to run the analysis both with and without the potentially bad data to get a sense of how much this might matter. In their Report, USPS both shows how the data are cleaned as well as demonstrates that the removal of outliers does not affect the key findings.

65. The USPS carefully detailed its data cleaning process in Library Reference USPS –RM2015-7/1 (filed December 11, 2014). Most removals were hard coded (i.e. "delete zip xxxxx") with a comment as to why. To give a sense of their data trimming, I provide some statistics given my reading of the USPS procedure. Out of 140,457 city carrier routes from USPS' Form 3999 data, the USPS dropped 82 routes because the routes occurred before 2009, 116 routes because Sunday activity was recorded, 313 routes because negative time was recorded for a delivery activity, 37 routes because delivery time exceeded 12 hours, and 42 routes because negative gross street time was recorded. In total, 508 routes were dropped, leaving 99.6 percent of the original data. Finally, the USPS decided to use data only from years 2012 and 2013, which reduced the number of routes to 112,972.

66. For USPS's collection volume field study, 72,178 route days were captured out of a possible 73,195. Among these 3,564 ZIP Code days captured, 51 were dropped due to missing data and 125 were imputed. USPS imputed ZIP Code days only when there were sufficient routes in the ZIP Code. All final 3,513 ZIP Code days were successfully merged with Form 3999 data and useable in USPS' final analysis. Finally, merging the above 3,513 ZIP Code days with the DOIS/CV data set resulted in 3,489 ZIP Code days. Two routes were removed because of incomplete data and 9 routes had

imputed data. I do not believe the Postal Service's removal of outliers is any cause for concern.

67. For USPS's package volume field study, 3,332 ZIP Code days recorded were merged with the 3999 dataset, resulting in 3,330 ZIP Code days (2 ZIP Code days in DOIS could not be matched with the 3999 dataset). Then, USPS Winsorized the data. Namely, they have time thresholds on both upper and lower end. For the lower end, the USPS analysts were worried about dropping low delivery times that were legitimate; hence, the analysts used a set of filters so that the data would be retained even if the low threshold was met. Nothing here strikes me as particularly odd or impactful.

## VI. CONCLUSION

68. For the reasons I discuss above, the Commission should reject both models – the National Form 3999 Model and Modified Proposal 13 Model – proposed by UPS in its June 8 comments. Additionally, UPS' criticisms of the Postal Service's proposal are without practical significance. Thus, I believe the Commission should feel comfortable approving the Postal Service's proposal in this proceeding.

## APPENDIX A – REVIEW OF DR. NEELS' MARCH 18 MODEL

69. In his June 8 supplemental report, Dr. Neels abandoned the multiplicative model that he proposed in his March 18 initial report. Thus, the main body of these comments discuss only the two models he proposes in his June 8 report, and not the earlier model. Dr. Neels' multiplicative model, however, is also flawed and should be rejected by the Commission as well. This appendix addresses two of the major problems with the model – improper aggregation of volume variables and use of a very unconventional, non-robust model specification –that are not indirectly addressed through the points made in the main body of my comments.

### A. The March 18 model does not include volume variables at an appropriate level of aggregation.

70. The first major flaw with Dr. Neels' March 18 proposal is that it aggregates all mail volume other than deviation parcels into a single variable. Neels Report (March 18, 2015) at 18-19. One might naturally expect (for operational reasons) that the link between city carrier street time and the various types of mail volume (e.g., delivery point sequenced letters, cased mail) could differ. This is certainly the Postal Service's established precedent. Petition of the United States Postal Service for the Initiation of a Proceeding to Consider Proposed Change in Analytical Principles (Proposal Thirteen) (filed December 11, 2014) at 4-5. To justify departing from this precedent, as Dr. Neels suggests, an econometrician should be expected to demonstrate that one cannot reject the null hypothesis that the relevant elasticities are all identical. In sharp contrast, the various null hypotheses (on either each volume pair, or perhaps better, one joint test on

all volumes – since this is explicitly the assumption being made by combining all non-parcel volumes) are largely rejected at conventional significance levels.

71. In its regular delivery panel regression, the USPS model regresses regular delivery time onto volume measures and controls:

$$\begin{aligned}\mathrm{DT} = {} & \beta_0 + \beta_1 DPS + \beta_2 CM + \beta_3 SEQ + \beta_4 FSS + \beta_5 CV \\ & + \beta_{1,2} DPS^2 + \beta_{2,2} CM^2 + \beta_{3,2} SEQ^2 + \beta_{4,2} FSS^2 + \beta_{5,2} CV^2 \\ & + \beta_{12} DPS \cdot CM + \beta_{13} DPS \cdot SEQ + \beta_{14} DPS \cdot FSS + \beta_{15} DPS \cdot CV \\ & + \beta_{23} CM \cdot SEQ + \beta_{24} CM \cdot FSS + \beta_{25} CM \cdot CV \\ & + \beta_{34} SEQ \cdot FSS + \beta_{35} SEQ \cdot CV + \beta_{45} FSS \cdot CV \\ & + \beta_{16} DPS \cdot DP + \beta_{26} CM \cdot DP + \beta_{36} SEQ \cdot DP + \beta_{46} FSS \cdot DP + \beta_{56} CV \cdot DP \\ & + \beta_6 DP + \beta_7 DM + \beta_8 MPDP + \beta_9 BR \\ & + \beta_{6,2} DP^2 + \beta_{7,2} DM^2 + \beta_{8,2} MPDP^2 + \beta_{9,2} BR^2\end{aligned}$$

where

$DT =$ Regular Delivery Time

$DPS =$ Delivery Point Sequenced Letters

$CM =$ Cased Mail

$SEQ =$ Sequenced Mail

$FSS =$ FSS Flats

$CV =$ Collection Volume

$DP =$ Delivery Points

$DM =$ Delivery Mode Indicator

$MPDP =$ Miles per Delivery Point

$BR =$ Proportion of Business Deliveries

(See USPS City Carrier Street Time Study report at p. 71.)

72. It is straightforward to test if the volume measures should be aggregated by restricting their coefficients to be identical, i.e., to test the following:

$$\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5$$
$$\beta_{1,2} = \beta_{2,2} = \beta_{3,2} = \beta_{4,2} = \beta_{5,2}$$
$$\beta_{16} = \beta_{26} = \beta_{36} = \beta_{46} = \beta_{56}$$

73. Using the USPS baseline panel regression as presented in its Report, the output of the Wald test on the joint restrictions that all volumes should be aggregated is as follows:[11]

| Test Statistic | Value | Df | Probability |
|---|---|---|---|
| Chi-square | 238.2664 | 12 | 0.0000 |

74. Performing a Wald test on the parameter restrictions (with a Chi-square value of 238.2664 and 12 degrees of freedom) reveals that we strongly reject the null hypothesis that the volume measures should be aggregated at conventional significance levels. To explore this further, I also performed pair-wise tests on the null hypotheses that the five volume measures' estimated coefficients are, pair-by-pair, identical. There are 10 pairs, in total.

[11] To keep things relatively simple, I dropped cross terms (which are included in USPS' original specification). That is, I impose that the cross terms for volumes are 0, i.e.

$$\beta_{12} = \beta_{13} = \beta_{14} = \beta_{15} = \beta_{23} = \beta_{24} = \beta_{25} = \beta_{34} = \beta_{35} = \beta_{45} = 0$$

I did not think it made sense to restrict the volume measures while allowing unrestricted cross terms. This is not, however, critical to the results.

## Results of Pair-Wise Tests

| Pair | | Difference in Marginal Costs | Standard Error | T-statistic | P-value | Is the difference statistically significant? |
|---|---|---|---|---|---|---|
| DPS | Cased Mail | -0.73 | 0.50 | -1.44 | 0.15 | No |
| DPS | Sequenced | -0.54 | 0.31 | -1.72 | 0.09 | No |
| DPS | FSS | -2.52 | 0.71 | -3.56 | 0.00 | Yes |
| DPS | Collection | -3.68 | 0.79 | -4.69 | 0.00 | Yes |
| Cased Mail | Sequenced | 0.19 | 0.50 | 0.38 | 0.70 | No |
| Cased Mail | FSS | -1.79 | 0.83 | -2.16 | 0.03 | Yes |
| Cased Mail | Collection | -2.96 | 0.91 | -3.25 | 0.00 | Yes |
| Sequenced | FSS | -1.98 | 0.70 | -2.82 | 0.00 | Yes |
| Sequenced | Collection | -3.15 | 0.85 | -3.70 | 0.00 | Yes |
| FSS | Collection | -0.54 | 0.98 | -0.55 | 0.58 | No |

75. Four of the pairs exhibit coefficient differences that are not statistically significant, indicating that the null hypothesis that the difference of the coefficients is 0 cannot be rejected. The remaining six pairs' differences are statistically significant. Dr. Neels reports only the four pairs (Table 1 of UPS CC.pdf) that are not statistically significant, and the code provided in his library reference fails to consider other combinations besides these selected four.

76. In sum, Dr. Neels argues for a specification designed to limit the proliferation of parameters. However, the consolidation of all non-parcel volume measures is unsupported by the data. As a result, his March 18 specification is mis-specified and should be rejected.

**B. The specification of Dr. Neels' March 18 model is not grounded in theory or common practice, nor are the conclusions one might draw from his specification robust.**

77. Dr. Neels estimates $\alpha$ , $\beta$ , $\delta$ , $\gamma_1$ , $\gamma_2$ , $\gamma_3$ from the cross-sectional specification. United Parcel Service Comments on Postal Service Proposal Thirteen Regarding City Carrier Street Time Costs, March 18, 2015, Exhibit A, at 24.

$$ST = \alpha \cdot DP^{\gamma_1} \left(\frac{SM}{DP}\right)^{\gamma_2} (NPV + \beta \cdot PV)^{\gamma_3}(1 + \delta \cdot DM) + \varepsilon$$

where

$ST$ = Total Street Time

$DP$ = Delivery Points

$SM$ = Street Miles

$NPV$ = Non Deviation Parcel Volume

$PV$ = Deviation Parcel Volume

$DM$ = Delivery Mode Indicator

78. This specification requires a non-linear solver to minimize the residuals from the estimation. The above specification converges after 62 iterations in Stata. The output is

| Source | SS | df | MS | | |
|---|---|---|---|---|---|
| | | | | Number of obs = | 292 |
| Model | 6203545.4 | 6 | 1033924.23 | R-squared = | 0.9761 |
| Residual | 151777.01 | 286 | 530.688845 | Adj R-squared = | 0.9756 |
| | | | | Root MSE = | 23.03668 |
| Total | 6355322.4 | 292 | 21764.8028 | Res. dev. = | 2654.657 |

| SH | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| /a | .0083416 | .0021084 | 3.96 | 0.000 | .0041917 | .0124915 |
| /gamma_1 | .4001799 | .0376053 | 10.64 | 0.000 | .3261616 | .4741982 |
| /gamma_2 | .0933336 | .0138145 | 6.76 | 0.000 | .0661427 | .1205245 |
| /beta | 11.46178 | 4.774624 | 2.40 | 0.017 | 2.063923 | 20.85964 |
| /gamma_3 | .5635369 | .0306074 | 18.41 | 0.000 | .5032926 | .6237813 |
| /delta | .3727051 | .0414221 | 9.00 | 0.000 | .2911742 | .454236 |

79. This *ad hoc* non-linear specification is novel, but Dr. Neels has provided no intellectual justification for it. To justify employing a nonconventional specification other than the transparent quadratic form used by USPS, one should have a very precise economic model that yields that form as an implication. Hence, demonstrating robustness of the key findings is a first order concern. The March 18 model, however, is not robust; modest variations in model specifications yield quite different results. For example, if we modify the specification to allow non-parcel and parcel volumes to have independent elasticities as follows:

$$ST = \alpha \cdot DP^{\gamma_1} \left(\frac{SM}{DP}\right)^{\gamma_2} NPV^{\gamma_3} (\beta \cdot PV)^{\gamma_4} (1 + \delta \cdot DM) + \varepsilon$$

then Stata will converge after 1 iteration with the following output

| Source | SS | df | MS | | |
|---|---|---|---|---|---|
| | | | | Number of obs = | 292 |
| Model | 2.328e-10 | 4 | 5.8208e-11 | R-squared = | 0.0000 |
| Residual | 1421997.2 | 287 | 4954.6941 | Adj R-squared = | -0.0139 |
| | | | | Root MSE = | 70.38959 |
| Total | 1421997.2 | 291 | 4886.58834 | Res. dev. = | 3307.979 |

| SH | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| /a | 129.9806 | . | . | . | . | . |
| /gamma_1 | 3.2e-313 | .0483595 | 0.00 | 1.000 | -.0951843 | .0951843 |
| /gamma_2 | 1.4e-313 | .0190423 | 0.00 | 1.000 | -.0374804 | .0374804 |
| /gamma1_3 | 6.6e-313 | .045003 | 0.00 | 1.000 | -.0885778 | .0885778 |
| /beta | 0 | (constrained) | | | | |
| /gamma_4 | 0 | 3.82e-14 | 0.00 | 1.000 | -7.52e-14 | 7.52e-14 |
| /delta | 4.3e-313 | .0425814 | 0.00 | 1.000 | -.0838114 | .0838114 |

80. That is, Stata's nonlinear solver immediately becomes stuck and is not able to reliably estimate the specification. Based upon my experience with nonlinear models, this implies that the optimization method may be trying to find a best fit over a very flat surface – simply put, it's hard for it to distinguish one solution from another. In such a

setting, I would be concerned about whether the optimization method has achieved the true best fit without additional, careful analysis.

81. Beyond this concern, it is extremely important to note that other reasonable versions of the specification that separate out a role for non-parcel and parcel volume suggest that parcel volume plays no role in predicting street hours, as evidenced by the 0 estimates for $\beta$ and $\gamma_4$.

82. I also dig a little deeper by considering the following simpler specifications:

$$ST = \alpha \cdot DP^{\gamma_1} \left(\frac{SM}{DP}\right)^{\gamma_2} PV^{\gamma_3}(1 + \delta \cdot DM) + \varepsilon$$

$$ST = \alpha \cdot PV^{\gamma} + \varepsilon$$

The only nonlinear specifications that do converge and have predictive power are specifications where parcel volume is not included, for example the specification

$$ST = \alpha \cdot DP^{\gamma_1} \left(\frac{SM}{DP}\right)^{\gamma_2} NPV^{\gamma_3}(1 + \delta \cdot DM) + \varepsilon$$

converges after 62 iterations, yielding:

| Source | SS | df | MS | | |
|---|---|---|---|---|---|
| | | | | Number of obs = | 292 |
| Model | 6199690.9 | 5 | 1239938.18 | R-squared = | 0.9755 |
| Residual | 155631.5 | 287 | 542.270027 | Adj R-squared = | 0.9751 |
| | | | | Root MSE = | 23.28669 |
| Total | 6355322.4 | 292 | 21764.8028 | Res. dev. = | 2661.98 |

| SH | Coef. | Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| /a | .0087225 | .0022201 | 3.93 | 0.000 | .0043528 | .0130923 |
| /gamma_1 | .396025 | .0380312 | 10.41 | 0.000 | .3211695 | .4708804 |
| /gamma_2 | .0930753 | .0139488 | 6.67 | 0.000 | .0656204 | .1205302 |
| /gamma_3 | .568414 | .0309307 | 18.38 | 0.000 | .5075343 | .6292937 |
| /delta | .3527984 | .0403921 | 8.73 | 0.000 | .2732961 | .4323007 |

83. Thus, Dr. Neels' nonlinear specification is well estimated only if (a) parcel volume is restricted to having the same elasticity as non-deviation parcel volume or (b) parcel volume is excluded. This is ***not*** a robust set of results from which one can confidently infer anything about the critical cost-volume relationships.

*Curriculum Vitae*

# CHRISTIAN T. LUNDBLAD

## OFFICE AND E-MAIL:

Christian T. Lundblad
Edward M. O'Herron Distinguished Scholar and
Professor of Finance
Ph.D. Program Coordinator
Kenan-Flagler Business School
The University of North Carolina at Chapel Hill
Campus Box 3490
McColl Building
Chapel Hill, NC 27599-3490
(919) 962-8441 (office)
(707) 371-7060 (fax)
Email: Christian_lundblad (at) unc.edu
Web address: http://public.kenan-flagler.unc.edu/faculty/lundblac

## ACADEMIC EXPERIENCE:

University of North Carolina, Chapel Hill, NC
- Edward M. O'Herron Distinguished Scholar and Professor of Finance — 2012 – present
- Associate Professor of Finance — 2008 – 2012
- Ph.D. Program Coordinator — 2010 – present
- Assistant Professor of Finance — 2006 – 2008

Indiana University, Bloomington, Indiana
- Assistant Professor of Finance — 2001 - 2006

Associate Editor, Journal of Finance — 2012 - present
Associate Editor, Financial Management — 2011 – present
Associate Editor, Journal of Banking and Finance — 2013 - present

## PROFESSIONAL EXPERIENCE:

Federal Reserve Board of Governors, Washington, DC
- Financial Economist — 2000-2001
- Division of International Finance
- Financial Markets Section
- Advised Board of Governors on International Financial Market Developments

Solutions Economics, Academic Affiliate, 2010-present

**EDUCATION**:

Ph.D., Financial Economics, Duke University, May 2000

M.A., Economics, Duke University, 1996

B.A., Economics and English Literature (with highest honors), Washington University in St. Louis, 1994

**VISITING POSITIONS:**

IGBS, Zagreb, Croatia, 2004, 2005, 2006

University of Naples Federico II, Naples, Italy, 2009

HEC Paris, 2009

Chinese University of Hong Kong, Hong Kong, China, 2010, 2013

Einaudi Institute for Economics and Finance, Rome, Italy, 2011

Banco Central de Reserva del Perú, Lima, Peru, 2012, 2013, 2014

Indian School of Business, Hyderabad, India, 2012, 2013, 2014

Tsinghua University (joint with UNC), Beijing, China, 2013, 2014

**ACADEMIC HONORS AND AWARDS**:

Teaching Excellence Award, MBA for Executives, 2009 & 2010

HEC-Paris Hedge Fund Research Grant, 2010-2011

Honorable mention, Smith-Breeden Distinguished Paper Award, 2005

3M Junior Faculty Research Grant 2004-2006

CIBER Research Grant, 2004-2006

Nominated for Kelley School of Business teaching award, 2005

Honorable mention, Barclay's Best Paper Prize, European Finance Association, 2004

Winner Barclay's Best Paper Prize, European Finance Association, 2001

Outstanding Teaching Award, Duke University, 1999

FLAS fellowship, 1996-1999

Phi Beta Kappa, 1994

**AREAS OF SPECIALIZATION**:

Primary Fields: Asset Pricing & Investments
International Finance & Emerging Markets
Secondary Fields: Econometrics & Statistics
Macroeconomics

**RESEARCH**:

PUBLICATIONS:

"Why do Term Structures in Different Countries Co-Move?," with Pab Jotikasthira and Anh Le, 2014, Journal of Financial Economics, forthcoming.

"Political Risk Spreads," with Geert Bekaert, Campbell R. Harvey, and Stephan Siegel, 2014. Journal of International Business Studies, forthcoming.

"The European Union, the Euro, and Equity Market Integration," with Geert Bekaert, Campbell R. Harvey, and Stephan Siegel, Journal of Financial Economics, 2013, 109, 583-603. (lead article).

"How Do Foreign Investors Impact Domestic Economic Activity? Evidence from China and India" with Tarun Ramadorai and Pab Jotikasthira, Journal of International Money and Finance, 2013, 39, 89-110.

"Asset Fire Sales and Purchases and the International Transmission of Financial Shocks," with Tarun Ramadorai and Pab Jotikasthira, Journal of Finance, 2012, 67, 2015-2050.

"Regulatory Pressure and Fire Sales in the Corporate Bond Market," with Andrew Ellul and Pab Jotikasthira, Journal of Financial Economics, 2011, 101, 596-620.

"What Segments Equity Markets?," with Geert Bekaert, Campbell R. Harvey, and Stephan Siegel, Review of Financial Studies, 2011, 24, 3841-3890 (lead article).

"Financial Openness and Productivity," with Geert Bekaert and Campbell R. Harvey, World Development, 2011, 39, 1-19 (lead article).

"Liquidity and Expected Returns: Lessons from Emerging Markets," with Geert Bekaert and Campbell R. Harvey, Review of Financial Studies, 2007, 20, 1783-1831.

"The Risk Return Tradeoff in the Long Run: 1836-2003," Journal of Financial Economics, 2007, 85, 123-150.

"Global Growth Opportunities and Market Integration," with Geert Bekaert, Campbell R. Harvey, and Stephan Siegel, Journal of Finance, 2007, 62, 1081-1137.

"Growth Volatility and Equity Market Liberalization," with Geert Bekaert and Campbell R. Harvey, Journal of International Money and Finance, 2006, 25, 370-403.

"Consumption, Dividends, and the Cross-Section of Equity Returns," with Ravi Bansal and Robert Dittmar, Journal of Finance, 2005, 60, 1639-1672.
Honorable mention, Smith-Breeden Distinguished Paper Award.

"Does Financial Liberalization Spur Growth," with Geert Bekaert and Campbell R. Harvey, Journal of Financial Economics, 2005, 77, 3-55 (lead article).

"Equity Market Liberalization in Emerging Markets," with Geert Bekaert and Campbell R. Harvey, Journal of Financial Research 2003 (Fall). Also published in The Federal Reserve Bank of St. Louis Review, July/August 2003, Volume 85, Number 4.

"Market Efficiency, Fundamental Values, and the Risk Premium in Global Equity Markets," with Ravi Bansal, Journal of Econometrics, 2002, 109, 195-237 (lead article).

"Emerging Equity Markets and Economic Development," with Geert Bekaert and Campbell R. Harvey. Journal of Development Economics, 2001, 66, 465-504.

OTHER PUBLICATIONS:

"The U.S. Economic Crisis: Root Causes and the Road to Recovery," with Greg Brown. Journal of Accountancy, 2009, 208, 42-49

"Behind the Boom: The Risks of Emerging Markets," Canadian Investment Review, 2007, 20, 22-26.

"Brazil in Crisis," with Campbell R. Harvey and Diego Valderrama, Emerging Markets Quarterly, 1999, Spring, 4-9.

BOOK CHAPTERS:

"Financial Openness and the Chinese Growth Experience," with Geert Bekaert and Campbell R. Harvey, in China's Financial Transition at a Crossroads, 2007, Charles Calomiris (ed.), Columbia University Press.

WORKING PAPERS:

"Is Historical Cost Accounting a Panacea? Market Stress, Incentive Distortions, and Gains Trading," with Andrew Ellul, Pab Jotikasthira, and Yihui Wang, 2013. (R&R)

"Mark-to-Market Accounting and Systemic Risk in the Financial Sector," with Andrew Ellul, Pab Jotikasthira, and Yihui Wang, 2013. (solicited for the 58 Economic Policy Panel for the journal Economic Policy, R&R)

"Stock Market Valuations Across U.S. States," with Geert Bekaert, Campbell R. Harvey, and Stephan Siegel, 2013.

"Endogenous Liquidity Supply," with Ravi Bansal and Wilbur John Coleman, 2011.

"Interpreting Risk Premia Across Size, Value, and Industry Portfolios," with Ravi Bansal and Robert Dittmar, 2003.

CONFERENCE PRESENTATIONS:

National Bureau of Economic Research
- "Consumption, Dividends, and the Cross-Section of Equity Returns," 2001
- "Interpreting Risk Premia across Size, Value and Industry Portfolios," 2003
- "Financial Openness and the Chinese Growth Experience," 2006
- "Liquidity and Financial Intermediation", 2009
- "Regulatory Pressure and Fire Sales in the Corporate Bond Market", 2009
- "Why do Term Structures in Different Countries Co-Move?," 2010

Western Finance Association:
- "Market Efficiency, Fundamental Values, and the Risk Premium in Global Equity Markets," 2000
- "Emerging Equity Markets and Economic Development," 2001
- "Consumption, Dividends, and the Cross-Section of Equity Returns," 2002
- "Growth Volatility and Equity Market Liberalization," 2003
- "Global Growth Opportunities and Market Integration," 2005
- "Why do Term Structures in Different Countries Co-Move?," 2011
- "Endogenous Liquidity Supply" 2011
- "Is Historical Cost Accounting a Panacea? Market Stress, Incentive Distortions, and Gains Trading," 2013

American Finance Association:
- "Emerging Equity Markets and Economic Development," 2001
- "Does Financial Liberalization Spur Growth?" 2002
- "Liquidity and Expected Returns: Lessons from Emerging Markets," 2004
- "What Segments Equity Markets?" 2009
- "Financial Openness and Productivity." 2010
- "Regulatory Pressure and Fire Sales in the Corporate Bond Market," 2011
- "Asset Fire Sales and Purchases and the International Transmission of Financial Shocks" 2011
- "Endogenous Liquidity Supply" 2011
- "The European Union, the Euro, and Equity Market Integration" 2012
- "Political Risk and International Valuation" 2013
- "Is Historical Cost Accounting a Panacea? Market Stress, Incentive Distortions, and Gains Trading," 2014

American Economic Association:
"What Segments Equity Markets?" 2007

Econometric Society:
"Growth Volatility and Equity Market Liberalization," 2003

Utah Winter Finance Meetings:
"Consumption, Dividends, and the Cross-Section of Equity Returns," 2002

European Finance Association:
"Emerging Equity Markets and Economic Development," 2000
"Does Financial Liberalization Spur Growth?" 2001
**Winner of Barclay's Global Best Paper Prize**
"Growth Volatility and Equity Market Liberalization," 2002
"Liquidity and Expected Returns: Lessons from Emerging Markets," 2003
"Interpreting Risk Premia Across Size, Value and Industry Portfolios," 2003
"Global Growth Opportunities and Market Integration," 2004
"The Risk Return Tradeoff in the Long-Run: 1836-2003," 2005
"What Segments Equity Markets?" 2007
"Asset Fire Sales and Purchases and the International Transmission of Financial Shocks" 2010
"Regulatory Pressure and Fire Sales in the Corporate Bond Market" 2010

World Bank Conference on Financial Globalization:
"Does Financial Liberalization Spur Growth?" 2002
"Global Growth Opportunities and Market Integration," 2005

World Bank Conference on Corporate Governance:
"What Segments Equity Markets?" 2007

Darden Emerging Markets Conference:
"Liquidity and Expected Returns: Lessons from Emerging Markets," 2005
"What Segments Equity Markets?" 2008
"Asset Fire Sales and Purchases and the International Transmission of Financial Shocks," 2010
"The European Union, the Euro, and Equity Market Integration," 2011
"Political Risk and International Valuation," 2012

University of Amsterdam Asset Pricing Retreat:
"Liquidity and Expected Returns: Lessons from Emerging Markets," 2005
"What Segments Equity Markets?," 2007

Washington University in St. Louis Asset Pricing Conference:
"What Segments Equity Markets?," 2007

St. Louis Federal Reserve Bank 27th Annual Economic Policy Conference, Finance and Real Economic Activity
"Equity Market Liberalization in Emerging Markets,"

Hotelling Triangle Econometrics Conference:
"Market Efficiency, Fundamental Values, and the Risk Premium in Global Equity Markets," 2001,

Conference on Financial Systems and Crises at the Yale School of Management:
"Does Financial Liberalization Spur Growth?" 2001

National Bureau of Economic Research - Inter-American Seminar of Economics:
"Emerging Equity Markets and Economic Development," 1999.

LARC meetings in Monterrey, Mexico
"Does Financial Liberalization Spur Growth?" 2001

INVITED RESEARCH PRESENTATIONS:
Harvard University (Economics), UCLA, University of North Carolina, University of Michigan, Duke University, University of Texas at Austin, Oxford University, Indiana University, University of Illinois, HEC-Paris, INSEAD, McGill University, University of Toronto, Hong Kong University of Science and Technology, Hong Kong University, Chinese University of Hong Kong, National University of Singapore, Singapore Management University, Nanyang Technical University of Singapore, University of Wisconsin, Indian School of Business, University of Utah, Emory University, University of California Irvine, Tilburg University, University of Amsterdam, Erasmus University, Board of Governors of the Federal Reserve, World Bank, College of William and Mary, Stockholm School of Economics, University of Lisbon, Babson College, Warwick Business School, University of Miami, Michigan State University, Simon Fraser University, Case Western University, St. Louis Federal Reserve Bank, Atlanta Federal Reserve Bank, University of Kansas, North Carolina State University.

INVITED/KEYNOTE CONFERENCE SPEECHES:
CEPR – Asset Pricing Seminar (invited focus session)
"Mis-Pricing and Cash Flow Risks," 2005
Emerging Markets Finance and Economics Meeting, Istanbul, Turkey
"Emerging Markets Liquidity," 2006
Brazilian Finance Conference, Sao Paulo, Brazil
"What Segments Equity Markets?" 2007

**TEACHING**:

MBA:
University of North Carolina:
EMBA (evening & weekend) Investments, 2007-present
EMBA (evening & weekend & One-MBA) Macroeconomics (core), 2008-present
EMBA (evening & weekend) Global Financial Markets, 2008-present
EMBA (joint with Tsinghua University) Global Financial Management, 2013-present
**Recipient of Teaching Excellence Award (2009, 2010)**

Executive Education:
University of North Carolina, Executive Development, 2009-present
(Manufacturing, Finance/Accounting, and U.S. military clients)

Tsinghua University, 2013, 2014

INSEAD/ILPSIE (Mumbai, India), 2013, 2014

Indian School of Business, MBA Global Macroeconomics, 2012, 2013. 2014

UNDERGRADUATE:
University of North Carolina:
Investments, 2006

Indiana University:
Intermediate Investments, 2001-2006
**Nominated for teaching award**

Duke University:
Financial Markets and Investments, 1999
**Recipient of teaching award for best graduate student teacher**

Ph.D.:
University of North Carolina:
Financial Economics, 2006-present

Indiana University:
Empirical Asset Pricing, 2003-2006

Duke University:
Mathematical Economics for Ph.D. students, 1999

SERVICE:
TO THE PROFESSION:

| | |
|---|---|
| Associate Editor, Journal of Finance | 2012 - present |
| Associate Editor, Financial Management | 2011 – present |
| Associate Editor, Journal of Banking and Finance | 2013 - present |

Ad-hoc Referee:
American Economic Review, Journal of Finance, Review of Financial Studies, Journal of Financial Economics, Journal of International Economics, Journal of Financial and Quantitative Analysis, Journal of Econometrics, Review of Economic Studies, Economic Journal, Journal of Empirical Finance, Journal of Financial Markets, Journal of Banking and Finance, Journal of Money, Credit, and Banking, Journal of Financial Intermediation, Journal of International Money and Finance, Journal of Applied Econometrics, Review of Finance, European Economic Review, Southern Economic Journal, Hong Kong - University Grant Committee

Program Co-Organizer:
Duke-UNC Asset Pricing Conference

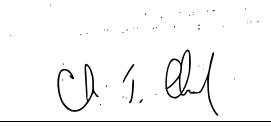
Program Committee:
Western Finance Association, 2008-present
Darden Emerging Markets Conference, 2008-present
SFS Cavalcade, 2011-present
Down Under Conference, 2011-present
European Finance Association, 2010-present
Napa Conference, 2011-present
Financial Management Association, 2006-present

TO THE SCHOOL:
University of North Carolina:
Ph.D. Area Coordinator, 2010-present
Member, FAC, 2008-2010
Alpha Challenge, 2009-2010

Ph.D Committee: Casey Dougal (2013), Isacco Piccioni (2012, chair), Matt Ringgenberg (2011), Wipawin Promboon (2009), Peter Groznik (2003), Pankaj Jain (2002), Sam Henkel.

## DECLARATION

I, Christian T. Lundblad, declare under penalty of perjury that the foregoing is true and correct. Executed on June 29, 2015.



Christian T. Lundblad